Final

Nulling Interferometry: Symmetry Requirements and Experimental Results

E. Serabyn

Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA

ABSTRACT

This paper provides a derivation from first principles of the stringent symmetry and stability requirements which deep stellar nulling demands, and also includes a brief status report on recent nulling results obtained with the Jet Propulsion Laboratory's fiber-coupled rotational-shearing interferometer. To date, the deepest transient nulls obtained (at red wavelengths) are 2×10^{-6} with a laser diode source, and 1.4×10^{-5} with a single-polarization thermal white-light source filtered to provide an 18% passband. In addition, both the laser and white light nulls have been stabilized to the 10^{-4} level. This visible wavelength laboratory nuller thus meets essentially all of the performance goals for the planned nulling experiment on board NASA's Space Interferometer Mission, with the sole exception of dual-polarization operation.

Keywords: Nulling interferometry, rotational shearing interferometer

1. INTRODUCTION

Although the high brightness contrast and small angular separation characterizing star-planet pairings has thus far prevented the *direct* detection of planets beyond our solar system, the novel technique of nulling interferometry shows great promise in overcoming these obstacles[1-10]. The basic premise of nulling interferometry is conceptually quite simple: interference of the light collected by a pair of telescopes to generate a deep destructive interference fringe at the stellar position, thus selectively dimming the star relative to its surroundings. Of course to be useful, the stellar cancellation must be both very deep and broadband. A null depth (the inverse of the rejection ratio) on the order of 10^{-6} is necessary to dim stars to the level of possible companion terrestrial planets at mid-infrared wavelengths, while bandwidths of $\Delta\lambda/\lambda \approx 100\%$ are needed to enable the detection of the extremely faint emission levels expected from terrestrial analogs. To effect such deep starlight cancellation, there is one simple experimental requirement: at zero optical path difference (OPD) between the incident stellar beams, the two arriving electric field vectors must subtract to high precision, i.e.,

$$\mathbf{E}_1 - \mathbf{E}_2 \approx 0.$$

However, meeting this requirement is decidedly non-trivial, as it calls for a very high degree of symmetry and stability in the optical system which collects and combines the starlight.

This paper has two main goals. The first is to derive the detailed symmetry and stability requirements which the above equation imposes on the optical system involved, and the second is to provide a brief status update on recent nulling experiments carried out at the Jet Propulsion Laboratory with a fiber-coupled rotational-shearing interferometer operating at visible wavelengths.

2. SYMMETRY AND STABILITY REQUIREMENTS

2.1. The case of plane monochromatic waves

As will be shown in the following, the basic requirement of precise electric field subtraction can be decomposed into a set of subsidiary symmetry and stability requirements. Because this field subtraction must occur for both incident polarization states, over a wide radiation passband, and across the full beam aperture, while also remaining stable in time, it is best to begin the analysis with the simplest case of an incident field which consists of plane monochromatic wavefronts. For this case, it is apparent that the two electric field vectors must be equal in amplitude and opposite in direction and phase. The relevant question then is how closely these matching conditions must be satisfied in order to achieve a given null depth, N, defined here as the ratio of the interferometer's transmission in the destructive and constructive states, viz., at the null fringe and at one of the two adjacent constructive peaks. This definition

was chosen for two reasons. First, because the constructive interference peak provides near-unity transmission for the stellar signal, N as defined is automatically normalized to the full stellar signal. And second, the alternative of normalizing by the signal at large OPD is unsatisfactory both because this approach would reference to only half of the maximum stellar signal, and because it is much more demanding experimentally, calling as it does for much larger, and hence more time-consuming, OPD changes.

For a plane monochromatic wavefront, the incident electric field vectors arriving at the two collectors are, of course, identical. After transmission down the two beam trains, the two electric fields can be described in terms of their amplitudes as

$$\mathbf{E}_1 = \mathrm{E}_{1x} e^{i\phi_{1x}} \hat{\mathbf{x}}_1 + \mathrm{E}_{1y} e^{i\phi_{1y}} \hat{\mathbf{y}}_1$$

and

$$\mathbf{E}_2 = \mathrm{E}_{2x} e^{i\phi_{2x}} \hat{\mathbf{x}}_2 + \mathrm{E}_{2y} e^{i\phi_{2y}} \hat{\mathbf{y}}_2$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are a pair of unit vectors defining the directions of the two orthogonal polarization states, and the subscripts 1 and 2 allow for a relative rotation of the fields in the two interferometer arms. The phase factors, ϕ_{ix} and ϕ_{iy}, with i = 1, 2, allow both for different path lengths in the two interferometer arms, and for the different phase delays likely to be experienced by the two polarization components in propagating down a single beam train.

At null, field subtraction occurs, while at the neighboring constructive peak the fields add. Summation and subtraction of the fields at a simple beamsplitter, or at a more complex, but equivalent, nulling beam combiner, gives

$$\mathbf{E}_1 \pm \mathbf{E}_2 = \frac{1}{\sqrt{2}} \left[\left(\mathrm{E}_{1x} e^{i\phi_{1x}} \hat{\mathbf{x}}_1 \pm \mathrm{E}_{2x} e^{i\phi_{2x}} \hat{\mathbf{x}}_2 \right) + \left(\mathrm{E}_{1y} e^{i\phi_{1y}} \hat{\mathbf{y}}_1 \pm \mathrm{E}_{2y} e^{i\phi_{2y}} \hat{\mathbf{y}}_2 \right) \right],$$

where the $1/\sqrt{2}$ factor reflects an ideal 50/50 intensity split in the beam combiner. Using the definition of intensity,

$$\mathrm{I} = \Re(\mathbf{E} \cdot \mathbf{E}^*/2),$$

where $\Re$ denotes the real part, the corresponding intensities for constructive (+) and destructive (-) interference are

$$\mathrm{I}_\pm = \frac{1}{4} \left[\mathrm{E}_{1x}^2 + \mathrm{E}_{2x}^2 \pm 2\mathrm{E}_{1x}\mathrm{E}_{2x} \cos(\phi_{1x} - \phi_{2x}) \hat{\mathbf{x}}_1 \cdot \hat{\mathbf{x}}_2 + \mathrm{E}_{1y}^2 + \mathrm{E}_{2y}^2 \pm 2\mathrm{E}_{1y}\mathrm{E}_{2y} \cos(\phi_{1y} - \phi_{2y}) \hat{\mathbf{y}}_1 \cdot \hat{\mathbf{y}}_2 \right]. \quad (1)$$

From this it is evident that the subtracted x-field intensity will be nonzero if, as mentioned earlier, $\phi_{1x} \neq \phi_{2x}$, or $\hat{\mathbf{x}}_1 \cdot \hat{\mathbf{x}}_2 \neq 1$, or $\mathrm{E}_{1x} \neq \mathrm{E}_{2x}$, with analogous conditions applying to the y-component of the field.

Introducing now the four component intensities, $\mathrm{I}_{ix} = \mathrm{E}_{ix}^2/2$ and $\mathrm{I}_{iy} = \mathrm{E}_{iy}^2/2$, with i=1, 2, and defining the relative phase delays for each polarization component as $\Delta\phi_x = \phi_{1x} - \phi_{2x}$ and $\Delta\phi_y = \phi_{1y} - \phi_{2y}$, as well as the relative polarization rotation angle $\alpha_{rot} = \cos^{-1}(\hat{\mathbf{x}}_1 \cdot \hat{\mathbf{x}}_2) = \cos^{-1}(\hat{\mathbf{y}}_1 \cdot \hat{\mathbf{y}}_2)$, equation (1) can be recast as

$$\mathrm{I}_\pm = \frac{1}{2} \left[\mathrm{I}_{1x} + \mathrm{I}_{2x} \pm 2\cos(\Delta\phi_x)\cos(\alpha_{rot}) \sqrt{\mathrm{I}_{1x}\mathrm{I}_{2x}} + \mathrm{I}_{1y} + \mathrm{I}_{2y} \pm 2\cos(\Delta\phi_y)\cos(\alpha_{rot}) \sqrt{I_{1y} I_{2y}} \right].$$

Furthermore, since deep nulling requires closely matched beam intensities, it is useful to introduce the mean intensities in each of the polarization states,

$$\langle \mathrm{I}_x \rangle = \frac{1}{2}(\mathrm{I}_{1x} + \mathrm{I}_{2x})$$

and

$$\langle \mathrm{I}_y \rangle = \frac{1}{2}(\mathrm{I}_{1y} + \mathrm{I}_{2y}),$$

and the fractional deviations from the mean intensities,

$$\delta I_x = \frac{I_{1x} - I_{2x}}{2\langle I_x \rangle}$$

and

$$\delta I_y = \frac{I_{1y} - I_{2y}}{2\langle I_y \rangle}$$

so that

$$I_\pm = \langle I_x \rangle \left[1 \pm \cos(\Delta\phi_x)\cos(\alpha_{rot})\sqrt{1-(\delta I_x)^2} \right] + \langle I_y \rangle \left[1 \pm \cos(\Delta\phi_y)\cos(\alpha_{rot})\sqrt{1-(\delta I_y)^2} \right]. \quad (2)$$

If $\Delta\phi_{x,y}$, α_{rot} and $\delta I_{x,y}$ are all small, the destructive term simplifies to

$$I_- = \frac{\langle I_x \rangle}{2}\left[(\Delta\phi_x)^2 + \alpha_{rot}^2 + (\delta I_x)^2\right] + \frac{\langle I_y \rangle}{2}\left[(\Delta\phi_y)^2 + \alpha_{rot}^2 + (\delta I_y)^2\right],$$

while the constructive maximum becomes

$$I_+ = 2\langle I_x \rangle + 2\langle I_y \rangle.$$

In the simplest case of a single polarization component, the null depth is then

$$N_{1pol} = \frac{I_-}{I_+} = \frac{1}{4}\left[(\Delta\phi_x)^2 + \alpha_{rot}^2 + (\delta I_x)^2\right], \quad (3)$$

while in the dual-polarization case with $\langle I_x \rangle \approx \langle I_y \rangle$, the result is

$$N_{2pol} = \frac{I_-}{I_+} = \frac{1}{8}\left[(\Delta\phi_x)^2 + (\Delta\phi_y)^2 + 2\alpha_{rot}^2 + (\delta I_x)^2 + (\delta I_y)^2\right]. \quad (4)$$

To recover the single-polarization result directly from equation (4), it is necessary to set $\Delta\phi_x = \Delta\phi_y$ and $\delta I_x = \delta I_y$.

The null depth is thus seen to be the linear sum of several independent light leakage (i.e. power) contributions, which arise in factors such as pathlength errors between the two interferometer arms, polarization rotation angle mismatch after transport down the two beam trains, and intensity imbalance between the two fields. The single polarization case is particularly simple, since only three light leakage terms contribute and the coefficients of all three terms are identical, so that the net null is given by

$$N = N_\phi + N_\alpha + N_I,$$

where these three terms correspond to those in equation (3).

In the dual polarization case, a revealing simplification results from the introduction of the mean (polarization-averaged) phase delay

$$\langle \Delta\phi \rangle = \frac{\Delta\phi_y + \Delta\phi_x}{2},$$

and the differential s-p phase delay (the difference between the s-p delays in the two interferometer arms, where the s-p delay in each arm is defined as the phase delay difference between the two [s and p] polarization components)

$$\Delta\phi_{s-p} = \Delta\phi_y - \Delta\phi_x.$$

With these substitutions, equation (4) becomes

$$N_{2pol} = \frac{1}{4}\left[(\Delta\phi)^2 + \frac{1}{4}(\Delta\phi_{s-p})^2 + \alpha_{rot}^2 + \frac{1}{2}(\delta I_x)^2 + \frac{1}{2}(\delta I_y)^2\right], \quad (5)$$

This form is more illuminating because it affords a more orthogonal separation of terms. The first term in equation (5) now clearly reflects the average phase (i.e. pathlength) error, which can be corrected by means of an optical delay line, while the second term is quite different, arising entirely in the differential s-p delay between the two beam trains, and so relates to asymmetries in the reflections and transmissions in the two beam trains. The third term is the polarization rotation angle, as in equation (3), while the last two terms are the intensity imbalances in the two polarization states. Equation (5) also reduces more transparently to the single polarization case of equation (3), in that the differential s-p delay term simply vanishes in the single-polarization case.

Two quite distinct sources of intensity mismatch can be envisioned: transmission asymmetries, which are static in character, and pointing-related time-dependent fluctuations of the coupling of the starlight into the single-mode spatial filter likely to be located in the nuller output focal plane[2]. Assuming that static intensity imbalances can be removed by an intensity control scheme of some sort, only the time-dependent pointing-induced intensity fluctuations remain. However, as such fluctuations affect both polarization components equivalently, the two intensity imbalance terms in equation (5) are equal in this case, yielding finally

$$\mathrm{N} = \frac{1}{4}\left[(\Delta\phi)^2 + \frac{1}{4}(\Delta\phi_{\mathrm{s-p}})^2 + \alpha_{\mathrm{rot}}^2 + (\delta\mathrm{I})^2\right], \tag{6}$$

where δI now refers to the fractional deviation from the mean total intensity, $\langle \mathrm{I} \rangle = \langle \mathrm{I_x} \rangle + \langle \mathrm{I_y} \rangle$. This equation is now in a form which is equally applicable to the single and dual polarization cases, since $\Delta\phi_{\mathrm{s-p}} = 0$ in the single polarization case, and the intensity term can be taken to refer to the total power present in either of the two cases. Equation (6) will be used as the starting point for the following generalization.

2.2. The Case of Polychromatic Plane Waves From a Source of Finite Extent

To summarize the development to this point, equations (5) and (6) give the instantaneous null depth for an incident monochromatic dual-polarization plane wave. However, real wavefronts will depart from this ideal case in several ways. First, the stellar spectrum is polychromatic. Second, the finite extent of a stellar source implies that in reality a superposition of plane waves with a small range of incidence angles is present. Third, after passage down the optical train the wavefronts will no longer be perfectly planar, but are likely to be somewhat aberrated on a variety of scales. Finally, the optical system is not likely to be completely stable, so that minute pathlength and pointing fluctuations will cause the null level to fluctuate. A generalization of the analysis is therefore required. In this section, the assumptions of monochromaticity and a point-like source are relaxed. The time dependencies caused by fluctuations will then be addressed in the following section. Due to space limitations, wavefront aberrations are not addressed here.

To begin, it is first important to recognize that the null depth N should more properly be written as a function of three variables, $\mathrm{N}(\theta, \lambda, \mathrm{t})$, where θ is the angular coordinate offset on the sky from, and in the direction normal to, the single-baseline interferometer's linear null fringe, λ is the operating wavelength, and t is the time. The experimentally measureable quantity is then the instantaneous null depth averaged over the normalized source distribution and passband,

$$\mathrm{N(t)} = \int\int \mathrm{N}(\theta, \lambda, \mathrm{t}) B(\Omega) S(\lambda)\ d\Omega d\lambda.$$

Here Ω is shorthand for the two-dimensional angular source coordinates, $B(\Omega)$ is the source's normalized spatial brightness distribution function and $S(\lambda)$ is the normalized detected spectrum (which accounts for both the incident spectrum and the instrumental transmission). For simplicity here, both $B(\Omega)$ and $S(\lambda)$ are assumed normalized to unity integrals. The main effect of both of these "weighting functions" is on the phase error term. Spectral issues may also impact the differential s-p delay and intensity terms, but likely only to second order, so these effects are neglected here. The integral of the phase term in equations (3), (5) and (6) is

$$\mathrm{N}_\phi(\mathrm{t}) = \int S(\lambda) d\lambda \left[\int \frac{(\Delta\phi(\theta, \lambda, \mathrm{t}))^2}{4} B(\Omega) d\Omega\right]. \tag{7}$$

Now $\Delta\phi$ is given by the sum of the on-axis phase delay, $\Delta\phi_{\mathrm{d}}(\lambda, \mathrm{t})$ (which fluctuates in time due to phase delay mismatches between the two interferometer arms), and the time-independent two-element interferometer fringe response pattern, $\phi_{\mathrm{fr}}(\theta, \lambda)$, i.e.,

$$\Delta\phi(\theta,\lambda,\mathrm{t}) = \Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t}) + \phi_{\mathrm{fr}}(\theta,\lambda).$$

Equation (7) then becomes

$$\mathrm{N}_{\phi}(\mathrm{t}) = \frac{1}{4}\int S(\lambda)d\lambda\left[\int\left[\left(\Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t}) + \phi_{\mathrm{fr}}(\theta,\lambda)\right)\right]^2 B(\Omega)d\Omega\right]. \tag{8}$$

For the case of a two-element nulling interferometer with a null fringe centered on a star, the fringe pattern has the familiar sinusoidal form

$$\phi_{\mathrm{fr}}(\theta,\lambda) = \frac{2\pi\mathrm{b}\sin(\theta)}{\lambda}, \tag{9}$$

where b is the baseline length. This function is antisymmetric in θ, while for a disk-like stellar source, $B(\Omega)$ is symmetric in θ, the perpendicular offset from the null fringe. Thus, in the expansion of equation (8),

$$\mathrm{N}_{\phi}(\mathrm{t}) = \frac{1}{4}\int S(\lambda)d\lambda\left[\int\left(\left(\Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t})\right)^2 + 2\Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t})\phi_{\mathrm{fr}}(\theta,\lambda) + \phi_{\mathrm{fr}}^2(\theta,\lambda)\right)B(\Omega)d\Omega\right],$$

the integral of the cross term is zero, so that

$$\mathrm{N}_{\phi}(\mathrm{t}) = \frac{1}{4}\int S(\lambda)d\lambda\left[\int\left(\Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t})\right)^2 B(\Omega)d\Omega\right] + \frac{1}{4}\int S(\lambda)d\lambda\left[\int\phi_{\mathrm{fr}}^2(\theta,\lambda)\,B(\Omega)d\Omega\right].$$

Since $\Delta\phi_{\mathrm{d}}(\lambda,t)$ is independent of the source parameters, this simplifies to

$$\mathrm{N}_{\phi}(\mathrm{t}) = \frac{1}{4}\int S(\lambda)\left(\Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t})\right)^2 d\lambda + \int S(\lambda)d\lambda\left[\frac{1}{4}\int\phi_{\mathrm{fr}}^2(\theta,\lambda)\,B(\Omega)d\Omega\right]. \tag{10}$$

The first term in this equation is the null depth for a point source as limited by a phase error between the two interferometer arms, while the second term is the null contribution due to the variable fringe transmission across the stellar disk. The latter integral can be done quite simply under the assumption of a disk-like source of small angular diameter θ_{dia}. The normalized source brightness distribution is then $B(\Omega) = (\pi\,\theta_{\mathrm{dia}}^2/4)^{-1}$ inside the stellar disk and zero outside, so with the fringe phase from equation (9), the term in square brackets in equation (10) becomes

$$\frac{1}{4}\int\phi_{\mathrm{fr}}^2(\theta,\lambda)\,B(\Omega)d\Omega = \frac{1}{\pi}\left(\frac{2\,\pi\mathrm{b}}{\theta_{\mathrm{dia}}\lambda}\right)^2\int_{\mathrm{star}}\theta^2\,\theta_{\mathrm{r}}\,d\theta_{\mathrm{r}}\,d\psi,$$

where the solid angle element $d\Omega$ has been replaced by its equivalent in terms of the radial and azimuthal angular coordinates θ_{r} and ψ. Recalling that θ is the angular offset from, and in the direction normal to, the null fringe, while θ_{r} is the radial coordinate offset from the center of the star, it is possible to write $\theta = \theta_{\mathrm{r}}\cos(\psi)$, so that

$$\frac{1}{4}\int\phi_{\mathrm{fr}}^2(\theta,\lambda)\,B(\Omega)d\Omega = \frac{1}{\pi}\left(\frac{2\,\pi\mathrm{b}}{\theta_{\mathrm{dia}}\lambda}\right)^2\int_0^{2\pi}\cos^2(\psi)\,d\psi\int_0^{\theta_{\mathrm{dia}}/2}\theta_{\mathrm{r}}^3\,d\theta_{\mathrm{r}},$$

which gives finally

$$\frac{1}{4}\int\phi_{\mathrm{fr}}^2(\theta,\lambda)\,B(\Omega)d\Omega = \frac{\pi^2}{16}\left(\frac{\theta_{\mathrm{dia}}}{\lambda/\mathrm{b}}\right)^2.$$

For this term to be small, a stellar diameter much smaller than the fringe spacing of λ/b is needed. Equation (10) then becomes

$$\mathrm{N}_{\phi}(\mathrm{t}) = \frac{1}{4}\int S(\lambda)\left(\Delta\phi_{\mathrm{d}}(\lambda,\mathrm{t})\right)^2 d\lambda + \frac{\left(\pi\,\mathrm{b}\,\theta_{\mathrm{dia}}\right)^2}{16}\int\frac{S(\lambda)}{\lambda^2}d\lambda. \tag{11}$$

The second term in this equation can be conservatively estimated simply by evaluating the integral at the shortest wavelength in the passband, $\lambda_{\rm sh}$, so that

$$\mathrm{N}_\phi(\mathrm{t}) = \frac{1}{4}\int S(\lambda)\left(\Delta\phi_{\rm d}(\lambda,\mathrm{t})\right)^2 d\lambda + \frac{\pi^2}{16}\left(\frac{\theta_{\rm dia}}{\lambda_{\rm sh}/b}\right)^2. \tag{12}$$

Evaluation of the first term in this equation now requires consideration of phase dispersion across the passband, which can be done by setting

$$\Delta\phi_{\rm d}(\lambda,\mathrm{t}) = \Delta\phi_{\rm c}(\mathrm{t}) + \Delta\phi_\lambda(\mathrm{t}),$$

where $\Delta\phi_{\rm c}(\mathrm{t})$ is the simple (time-dependent) geometric phase delay at band center, while the $\Delta\phi_\lambda(\mathrm{t})$ term accounts for phase dispersion across the passband. In vacuum, the phase dispersion term can be assumed to be time-independent, but for the case of ground-based interferometers with open-air delay lines, the time-variable amounts of H_2O vapor present in the two beam trains can lead to the time-dependence of this term also. Expanding the integral in equation (12) then gives

$$\int S(\lambda)\left(\Delta\phi_{\rm d}(\lambda,\mathrm{t})\right)^2 d\lambda = \int S(\lambda)\left[(\Delta\phi_{\rm c}(\mathrm{t}))^2 + 2\Delta\phi_{\rm c}(\mathrm{t})\Delta\phi_\lambda(\mathrm{t}) + (\Delta\phi_\lambda(\mathrm{t}))^2\right] d\lambda.$$

Assuming for simplicity a linear phase gradient with wavelength, the middle term would disappear if $S(\lambda)$ were a perfectly flat function. However, even if $S(\lambda)$ is not flat, the third term will dominate the second due to the combination of its positive definite nature and the zero-mean nature of $\Delta\phi_{\rm c}(\mathrm{t})$ under the action of a control loop (next section). It is therefore safe to neglect the middle term, yielding,

$$\int S(\lambda)\left(\Delta\phi_{\rm d}(\lambda,\mathrm{t})\right)^2 d\lambda = \left(\Delta\phi_{\rm c}(\mathrm{t})\right)^2 + \int S(\lambda)(\Delta\phi_\lambda(\mathrm{t}))^2 d\lambda,$$

since $\Delta\phi_{\rm c}(\mathrm{t})$ is independent of λ and $S(\lambda)$ has an integral of unity. The first term here is the simple geometric path delay error at the passband center, which is a function only of time, while the second term is the spectrally-weighted mean square dispersion across the passband,

$$\langle(\Delta\phi_\lambda(\mathrm{t}))^2\rangle = \int S(\lambda)(\Delta\phi_\lambda(\mathrm{t}))^2 d\lambda.$$

The net phase error term in equation (12) thus becomes

$$\mathrm{N}_\phi(\mathrm{t}) = \frac{\left(\Delta\phi_{\rm c}(\mathrm{t})\right)^2}{4} + \frac{\langle(\Delta\phi_\lambda(\mathrm{t}))^2\rangle}{4} + \frac{\pi^2}{16}\left(\frac{\theta_{\rm dia}}{\lambda_{\rm sh}/\mathrm{b}}\right)^2. \tag{13}$$

The three phase-related terms here are due to first, the polarization-averaged instantaneous phase error at the passband center (which is related to delay-line positioning errors), second, the phase dispersion across the passband (which is related to beamsplitter coatings and other dielectrics in the beam train, including H_2O vapor), and third, the leakage due to the finite extent of the star (which can be altered either by observing a more distant star, or by changing the observing wavelength or the baseline length).

Finally, inserting $\mathrm{N}_\phi(\mathrm{t})$ into equation (6) in place of the simple $(\Delta\phi)^2/4$ phase term, and now also explicitly allowing for the time-dependence of the intensity term, the instantaneous null depth in the more general case of polychromatic plane waves from a disk-like source becomes

$$\mathrm{N}(\mathrm{t}) = \frac{1}{4}\left[\left(\Delta\phi_{\rm c}(\mathrm{t})\right)^2 + \langle(\Delta\phi_\lambda(\mathrm{t}))^2\rangle + \frac{\pi^2}{4}\left(\frac{\theta_{\rm dia}}{\lambda_{\rm sh}/\mathrm{b}}\right)^2 + \frac{1}{4}\left(\Delta\phi_{\rm s-p}\right)^2 + \alpha_{\rm rot}^2 + \left(\delta\mathrm{I}(\mathrm{t})\right)^2\right]. \tag{14}$$

2.3. Fluctuations

Of the six terms in equation (14), to first order only the initial phase-delay term and the last intensity term are expected to show significant time-dependent fluctuations in vacuum, the former due to minute path delay fluctuations, and the latter due to fluctuations in telescope pointing, which lead to fluctuations in the coupling of the starlight into the requisite single-mode spatial filter. For simplicity, water vapor column fluctuations will be neglected hereafter, so that the dispersion term will be assumed constant.

To calculate the magnitude of the fluctuations in the null level, it is necessary to address the mean and the variance of the null. The time-averaged null,

$$\overline{\mathrm{N}} = \frac{1}{\mathrm{T}} \int \mathrm{N(t)}\ d\mathrm{t},$$

with T the integration time, is given by

$$\overline{\mathrm{N}} = \frac{1}{4}\left[\overline{(\Delta\phi_{\mathrm{c}}(\mathrm{t}))^2} + \langle(\Delta\phi_\lambda)^2\rangle + \frac{\pi^2}{4}\left(\frac{\theta_{\mathrm{dia}}}{\lambda_{\mathrm{sh}}/b}\right)^2 + \frac{1}{4}(\Delta\phi_{\mathrm{s-p}})^2 + \alpha_{\mathrm{rot}}^2 + \overline{(\delta\mathrm{I(t)})^2}\ \right], \tag{15}$$

If control loops are in place to stabilize both of the fluctuating quantities, OPD and pointing, at zero mean, i.e.,

$$\overline{\Delta\phi_{\mathrm{c}}(\mathrm{t})} = 0,$$

and

$$\overline{\delta\mathrm{I(t)}} = 0,$$

the first term in equation (15) is then the variance of the phase error,

$$\sigma_\phi^2 = \overline{(\Delta\phi_{\mathrm{c}}(\mathrm{t}))^2}$$

while the last term is the variance of the fractional intensity deviations,

$$\sigma_{\mathrm{I}}^2 = \overline{(\delta\mathrm{I(t)})^2}.$$

The time-averaged null in the presence of active OPD matching and intensity balancing is thus

$$\overline{\mathrm{N}} = \frac{1}{4}\left[\sigma_\phi^2 + \langle(\Delta\phi_\lambda)^2\rangle + \frac{\pi^2}{4}\left(\frac{\theta_{\mathrm{dia}}}{\lambda_{\mathrm{sh}}/b}\right)^2 + \frac{1}{4}(\Delta\phi_{\mathrm{s-p}})^2 + \alpha_{\mathrm{rot}}^2 + \sigma_{\mathrm{I}}^2 \right], \tag{16}$$

The variance of the null, given by

$$\sigma_{\mathrm{N}}^2 = \frac{1}{\mathrm{T}} \int (\mathrm{N(t)} - \overline{\mathrm{N}})^2\ d\mathrm{t}\,,$$

is then

$$\sigma_{\mathrm{N}}^2 = \frac{1}{16\,\mathrm{T}} \int \left[\left((\Delta\phi_{\mathrm{c}}(\mathrm{t}))^2 - \sigma_\phi^2\right) + \left((\delta\mathrm{I(t)})^2 - \sigma_{\mathrm{I}}^2\right) \right]^2\ d\mathrm{t}. \tag{17}$$

Given that pointing and OPD fluctuations are entirely uncorrelated, it is safe to assume that

$$\overline{(\Delta\phi_{\mathrm{c}}(\mathrm{t}))^2(\delta\mathrm{I(t)})^2} = \sigma_\phi^2\sigma_{\mathrm{I}}^2,$$

so that the cross term in the outer square can be neglected. Furthermore, using the fact that for a zero-mean Gaussian random process in ξ with a standard deviation of σ_ξ the expected value of ξ^4 is given by

$$\overline{\xi^4} = 3\sigma_\xi^4 \ ,$$

equation (17) reduces to

$$\sigma_{\mathrm{N}}^2 = \frac{1}{16}(\ 2\sigma_\phi^4 + 2\sigma_{\mathrm{I}}^4) \ ,$$

so that the root-mean-square (rms) fluctuation of the null level is

$$\sigma_{\mathrm{N}} = \sqrt{\frac{\sigma_\phi^4 + \sigma_{\mathrm{I}}^4}{8}} \ . \tag{18}$$

A fluctuation of the instantaneous null to a level x standard deviations above the mean null level, where x is an arbitrary number, is then

$$\mathrm{N}_{\mathrm{x}\sigma_{\mathrm{N}}} = \overline{\mathrm{N}} + \mathrm{x}\sigma_{\mathrm{N}} = \overline{\mathrm{N}} + \mathrm{x}\sqrt{\frac{\sigma_\phi^4 + \sigma_{\mathrm{I}}^4}{8}} \ . \tag{19}$$

Here $\overline{N}$ is given by equation (16). In the case where the mean null is in fact determined entirely by fluctuations in the two time-variable quantities, equation (19) simplifies to

$$\mathrm{N}_{\mathrm{x}\sigma_{\mathrm{N}}} = \frac{\sigma_\phi^2 + \sigma_{\mathrm{I}}^2}{4} + \mathrm{x}\sqrt{\frac{\sigma_\phi^4 + \sigma_{\mathrm{I}}^4}{8}} \ , \tag{20}$$

In the more restrictive case that pathlength fluctuations alone dominate the null, this becomes

$$\mathrm{N}_{\mathrm{x}\sigma_{\mathrm{N}}} = \sigma_\phi^2 \frac{(1 + \mathrm{x}\sqrt{2})}{4} \ , \tag{21}$$

while if on the other hand intensity fluctuations alone dominate, the analogous equation is

$$\mathrm{N}_{\mathrm{x}\sigma_{\mathrm{N}}} = \sigma_{\mathrm{I}}^2 \frac{(1 + \mathrm{x}\sqrt{2})}{4} \ . \tag{22}$$

To understand the meaning of these last two results, it is useful to consider one of these two limiting cases further. For dominant pathlength fluctuations, the instantaneous pathlength phase error to which the $\mathrm{x}\sigma_{\mathrm{N}}$ null level corresponds (from equation 14) is simply $\phi(\mathrm{x}\sigma_{\mathrm{N}}) = 2\sqrt{\mathrm{N}_{\mathrm{x}\sigma_{\mathrm{N}}}}$, or

$$\phi(\mathrm{x}\sigma_{\mathrm{N}}) = \sigma_\phi \sqrt{1 + \mathrm{x}\sqrt{2}}. \tag{23}$$

This equation results from the quadratic relationship between the null depth and the phase error, and reflects the fact that in this limit the phase fluctuations themselves actually set both the mean null level and the rms fluctuation level. In fact, equation (23) implies that null levels 1-σ and 2-σ above the mean null occur for phase errors of $1.55\sigma_\phi$ and $1.96\sigma_\phi$, respectively. For Gaussian phase fluctuations, null leakages at or above these 1-σ and 2-σ null levels then occur only 12% and 5% of the time. Thus, because of the squaring process, the null leakage is actually somewhat more stable than the phase, implying that fluctuations in the null leakage 1-σ above the average null are already quite rare. Thus in the following, a 2-σ (95% confidence inteval) null stability criterion is used.

2.4. Summary of Results

Equations (14), (16) and (18) thus provide a complete description of the nulling performance to be expected in the general case of plane polychromatic waves from a small disk-like source. Specifically these equations provide the instantaneous, mean and rms null levels. These equations can thus be used to set an experimental error budget on the null leakage contributions. If each instantaneous individual contribution to the light leakage at null is required to be less than some maximum allowable null contribution, N_c (not necessarily equal for each contribution), equation (14) leads to six constraints, the first four of which are essentially identical in form:

$$\Delta\phi_c < 2\sqrt{N_c} \tag{24a}$$

$$\delta I < 2\sqrt{N_c} \tag{24b}$$

$$\alpha_{rot} < 2\sqrt{N_c} \tag{24c}$$

$$\sqrt{\langle(\Delta\phi_\lambda)^2\rangle} < 2\sqrt{N_c} \tag{24d}$$

$$\Delta\phi_{s-p} < 4\sqrt{N_c} \tag{24e}$$

$$\theta_{dia} < \frac{4}{\pi}\frac{\lambda_{sh}}{b}\sqrt{N_c}\ . \tag{24f}$$

Since the sum of these six error terms must be smaller than the desired net null depth, N_{net}, a safe margin is then present if each contribution is roughly $N_c \approx N_{net}/10$.

However, the first two of these relations should more properly be given in terms of rms fluctuation levels. Equations (21) and (22) then imply that relations (24a) and (24b) should be replaced by

$$\sigma_\phi < \frac{2\sqrt{N_c}}{\sqrt{1+x\sqrt{2}}} \tag{24a'}$$

$$\sigma_I < \frac{2\sqrt{N_c}}{\sqrt{1+x\sqrt{2}}}\ . \tag{24b'}$$

If x is set equal to 2 (corresponding to a positive 2-σ null fluctuation), these last two relationships can be approximated quite closely by

$$\sigma_\phi < \sqrt{N_c} \tag{24a''}$$

$$\sigma_I < \sqrt{N_c}\ , \tag{24b''}$$

both of which are a factor of two more stringent than equations (24a) and (24b).

Finally, it is important to note that although the set of six equations given here (24a''-b'' and 24c-f) rigorously describes the nulling error budget in the case considered, two issues nevertheless remain absent from the treatment. These are the effects of wavefront aberrations and photon noise. Both of these factors are certain to affect nulling performance, but time and space constraints do not permit treatment of these topics here.

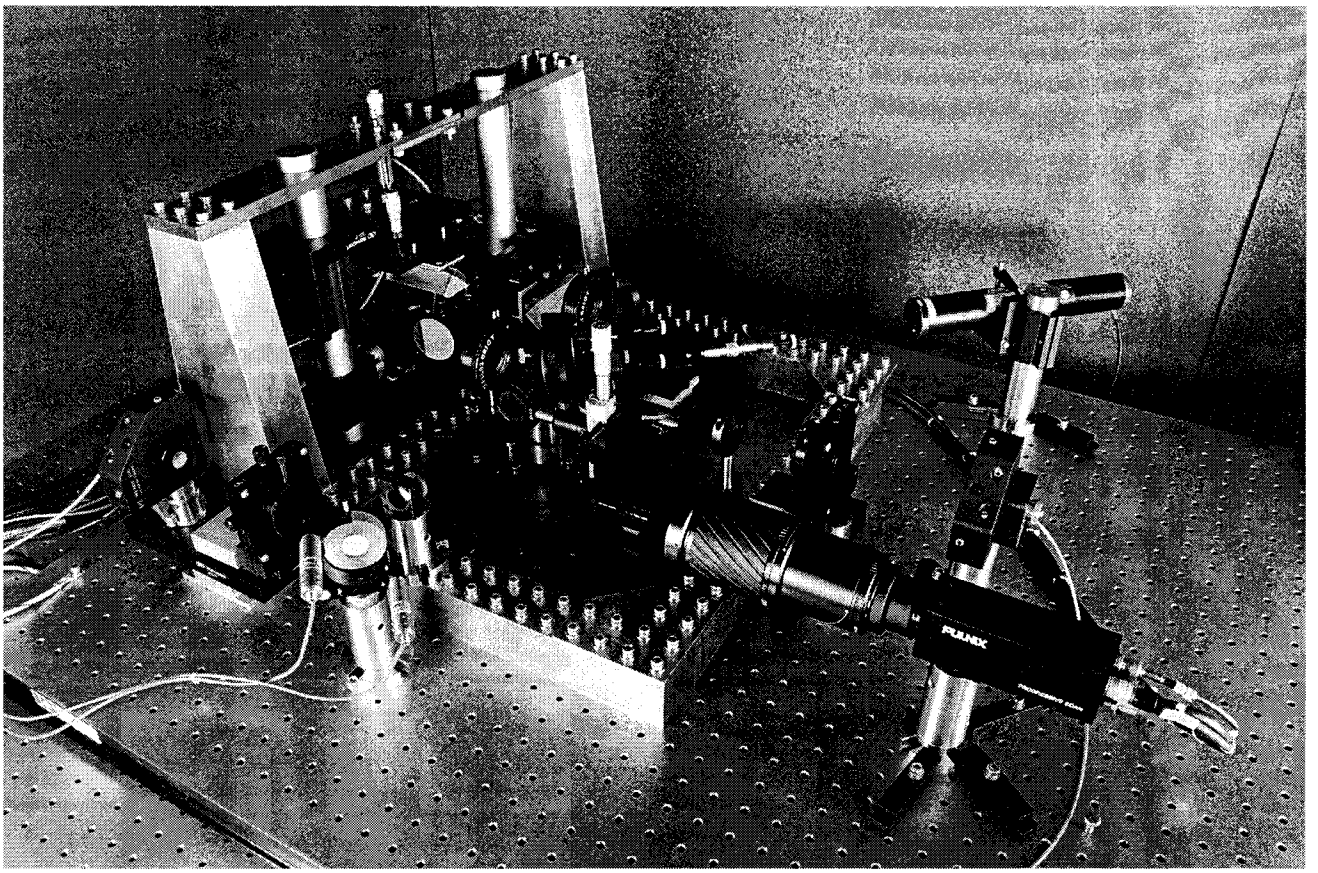

Figure 1. Fiber-coupled rotational-shearing interferometer[5,9,10] used in the JPL nulling experiments.

3. EXPERIMENTAL NULLING RESULTS

Finally, a brief status update on the nulling experiments being carried out at JPL is in order. The JPL laboratory nuller is a fiber coupled rotational shearing interferometer operating at visible wavelengths[5,9], and it had earlier demonstrated deep nulling of visible laser light[5]. After further upgrades to the experimental setup, stable white light nulls of better than a part in 10^4 are now obtained regularly[10] with single-polarization, 18% bandwidth white light. An example of such a stabilized white light null is shown in Figure 2, and an overview of the current experimental status, in terms of best nulls vs. bandwidth, is given in Figure 3. With this level of demonstrated performance, this laboratory nuller has thus met essentially all of the performance goals of the nulling experiment planned for NASA's Space Interferometer Mission (SIM), with the single exception of dual-polarization operation.

ACKNOWLEDGMENTS

The experimental results mentioned herein were obtained with the adroit assistance of J. K. Wallace and G. J. Hardy, and in the latter stages, S. Martin. Thanks are also due to G. J. Hardy and C. Koresko for meticulous comments on the manuscript, and to M. Shao and M.M. Colavita for helpful discussions. Funding was provided by NASA through the Keck and SIM interferometer projects. This work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

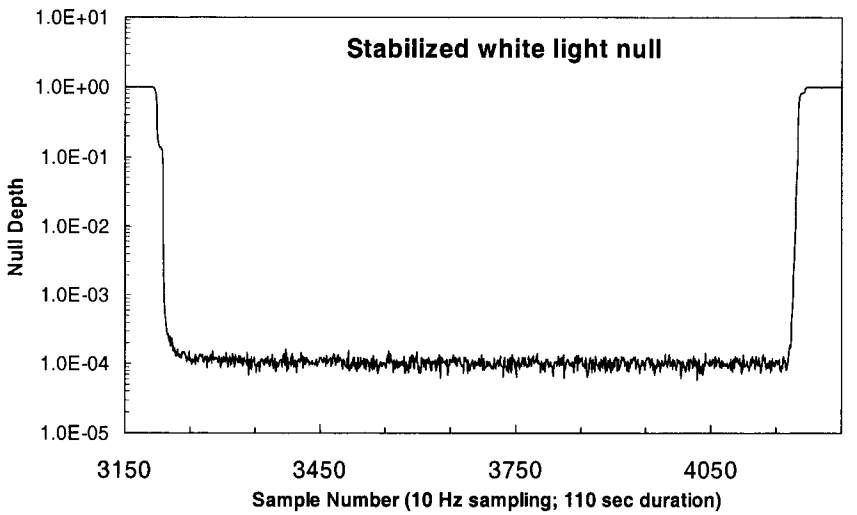

Figure 2. Stabilized white-light null for 18% bandwidth, single-polarization thermal light centered at 650 nm. The plotted time scan of the nuller output begins with the nuller locked on a constructive interference fringe by means of a position-dither feedback loop. The phase of the lock loop is then flipped by 180 degrees, and the nuller quickly locks onto the null fringe, rejecting light to the 10^{-4} level. Finally, the lock loop phase is flipped a second time, returning the nuller to the constructive fringe.

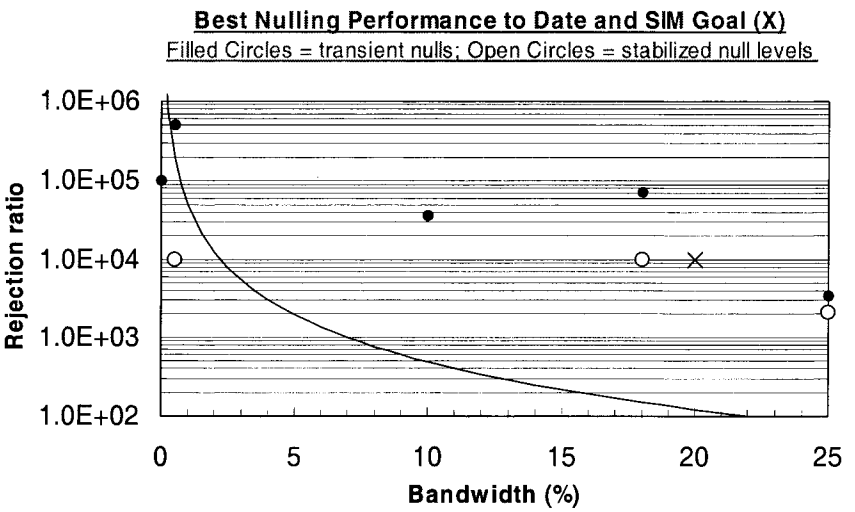

Figure 3. Status of the JPL nulling experiments as of late March 2000. The plot shows both the best transient and stabilized rejection ratios obtained vs. radiation bandwidth. The curve shows the best rejection which could be obtained with a standard laboratory Michelson interferometer.

REFERENCES

1. Bracewell, R.N. 1978, Nature 274, 780
2. Shao, M. & Colavita, M.M. 1992, Ann. Rev. Astron. Astrophys. 30, 457
3. Woolf, N. & Angel, J.R. 1998, Ann. Rev. Astron. Astrophys. 36, 507
4. Hinz, P.M. *et al.* 1998, Nature 395, 251
5. Serabyn, E., Wallace, J.K., Hardy, G.J., Schmidtlin, E.G.H. & Nguyen, H.T. 1999, Applied Optics 38, 7128
6. Beichman, C.A., Woolf, N.J. & Lindensmith, C.A. 1999, Terrestrial Planet Finder, JPL publ. 99-3
7. Leger, A., Mariotti, J., Mennesson, B., Ollivier, M., Puget, J., Rouan, D. & Schneider, J. 1996, Icarus, 123, 249
8. Mennesson, B. & Mariotti, J.M. 1997, Icarus 128, 202
9. Serabyn, E. 2000, in Darwin and Astronomy: the Infrared Space Interferometer, ESA publ. SP-451, in press
10. Wallace, J.K., Hardy, G.J., & Serabyn, E. 2000, in prep.